

**Note to Readers:** *EHP* strives to ensure that all journal content is accessible to all readers. However, some figures and Supplemental Material published in *EHP* articles may not conform to 508 standards due to the complexity of the information being presented. If you need assistance accessing journal content, please contact [ehp508@niehs.nih.gov](mailto:ehp508@niehs.nih.gov). Our staff will work with you to assess and meet your accessibility needs within 3 working days.

## **Supplemental Material**

### **Particulate Air Pollution Exposure and Expression of Viral and Human MicroRNAs in Blood: The Beijing Truck Driver Air Pollution Study**

Lifang Hou, Jitendra Barupal, Wei Zhang, Yinan Zheng, Lei Liu, Xiao Zhang, Chang Dou, John P. McCracken, Anaité Díaz, Valeria Motta, Marco Sanchez-Guerra, Katherine Rose Wolf, Pier Alberto Bertazzi, Joel D. Schwartz, Sheng Wang, and Andrea A. Baccarelli

#### **Table of Contents**

**Figure S1.** Enrichment factors (EFs) of the elemental components of PM<sub>2.5</sub> in truck drivers (TDs) relative to office workers (OW). EFs were computed as follows:  $EF(x) = \frac{[x_{(TD)}/PM_{2.5(TD)}]}{[x_{(OW)}/PM_{2.5(OW)}]}$ ; where  $[x_{(TD)}/PM_{2.5(TD)}]$  is the ratio of the mean concentration of element x over the mean PM<sub>2.5</sub> mass in truck drivers; and  $[x_{(OW)}/PM_{2.5(OW)}]$  is the ratio of the mean concentration of the same element x over the mean PM<sub>2.5</sub> mass in office workers. EF >1 indicates the elemental component concentrations in truck driver exceed those expected in office workers.

**Figure S2.** Cross-platform correlation across 20 randomly selected independent samples. Confirmation of the nCounter miRNA data was evaluated in 20 randomly selected samples using the TaqMan OpenArray microRNA expression profiling with the QuantStudio 12K Flex Real-Time PCR System. The two independent platforms demonstrated comparable miRNA profiles in these random samples.

**Figure S3.** Overview of the bioinformatics and biostatistics analysis. The raw miRNA expression data from 228 samples was processed using the NanoStringNorm R package. Target mRNAs of miRNAs associated with short-term changes in EC exposure were selected from three

databases (miRTarBase, TarBase, and miRBase). For genes selection in miRBase, two cutoffs,  $10^{-6}$  and  $10^{-8}$ , were chosen. Genes selected in miRBase with a cutoff of  $10^{-8}$  along with genes from experimental databases were used for Reactome functional networks analyses. FDR, false discovery rate.

**Figure S4.** Overview of the sample relations based on the PCA plot from 240 observations. A total of 12 observations (all were different participants) in the elongated circle were considered to be obvious outliers based on the PCA plot. NanoString plate numbers were labeled for these 12 samples.

**Figure S5.** Technical robustness of NanoString plates. Plate numbers are shown on the X-axis. Expression values of positive controls ( $\log_2$ ) are shown on the Y-axis. Each bar represents a plate. Plates were technically robust ( $P$ -value = 0.9951, one-way ANOVA).

**Figure S6.** Plate effects on miRNA expression based on PCA plot using pre-processed data. Numbers in the plot represent plate numbers. Negligible plate effects visually observed. Though samples were not randomized by examination day, distinct pattern of miRNA profiles due to plate design were not observed.

**Figure S7.** Venn diagram showing number of enriched pathways by KEGG and GO (Biological Process) analysis. miRBase assigns  $P$ -values to individual miRNA-3'-UTR target binding site, so different set of genes can be predicated with a miRBase threshold. Circle color represents miRBase threshold ( $10^{-6}$  in blue and  $10^{-8}$  in green). KEGG analysis shows that in the pooled analysis of both groups, only one pathway was found at both thresholds (A). In office workers, most of the pathways (17) were found at both thresholds (B). In truck drivers, 17 out of 19 total pathways were found at both thresholds (C). GO analysis shows that no GO term was robust in both group analyses (D). In office workers, 23 of the 31 total GO terms were found at both thresholds. In truck drivers, 13 of 20 total GO terms were found at both thresholds (F).

**Figure S8.** Functional analysis using Gene Ontology (GO) biological process terms. GO analysis was performed using combined target genes of all miRNAs with significant changes in each analysis group. Gene sets containing a minimum of eight genes at  $FDR < 1\%$  were considered significant. miRBase assigns  $P$ -values to individual miRNA-3'-UTR target binding site, so different set of genes can be predicated with a miRBase threshold. Two threshold cutoffs (i.e.  $10^{-6}$  and  $10^{-8}$ ) in miRBase database are shown on X-axis. Color of a tile represents enrichment of a GO term; Green= YES, Red= NO. GO term classification which is based on title of a GO term were shown on the right side. In the pooled analysis of both groups, 3 out of 5 total biological processes were related to biological regulation. In office workers, more than 35% of biological processes were related to immune response, and in truck drivers, 50% of biological processes were related to cell proliferation and differentiation. Two biological processes were common in all analysis groups. Seven biological processes were common between office workers and truck drivers.

**Table S1.** Percent change in miRNA expression level with one  $\mu\text{g}/\text{m}^3$  increase in Elemental Carbon (EC) at FDR <20%.

**Table S2.** Significance of interaction between the exposure (short term exposure to EC) and occupational groups for miRNAs associated with EC by group analysis at FDR<10%.

**Table S3.** Percent change in miRNA expression level with one-unit increase in elemental carbon (EC) (non-smokers only).

**Table S4.** Change in miRNA expression level (log2) with one unit increase in  $\text{PM}_{10}$ .

**Table S5.** Percent change in miRNA expression level with one-unit increase in  $\text{PM}_{2.5}$ .

**Table S6.** KEGG pathway enrichment analysis for targets of significant EC-associated miRNAs (for miRBase gene targets at  $P$ -value  $<10^{-8}$ ).

**Table S7.** Gene Ontology (biological term) enrichment analysis (FDR< 1%) for significant EC-associated miRNAs (miRBase gene targets at  $P$ -value  $<10^{-8}$ ).